Contingency Theory and Moderated Regression Analysis: the Effect of ...

Gelderman, M.

2000

document version
Early version, also known as pre-print

ıculteit der Economische Wetenschappen en Econometrie

SERIE RESEARCH MEMORANDA

Contingency Theory and Moderated Regression Analysis: the Effect of Measurement Level, Measurement Error and Non-Linear Relations

Maarten Celderman

Research Memorandum 2000-32 October 2000

vrije Universiteit amsterdam

Note:

Contingency Theory and Moderated Regression Analysis: the Effect of Measurement Level, Measurement Error and Non-linear Relations*

Maarten Gelderman

mgelderman@planet.nl

Vrije Universiteit Amsterdam

Department of Business and Economics

De Boelelaan 1105-kr 2E62

NL 1081 HV Amsterdam

The Netherlands

+31 20 444 6005 (fax)

+31 20 444 6040 (phone)

+31 20 623 75 80 (home phone)

October 23, 2000

*I wish to thank Eelke Wiersma for comments on an earlier version of this paper.

Abstract

Moderated regression analysis is generally accepted as the most appropriate way to asses the viability of contingency models. This paper discusses some problems associated with the application of this technique. If interval data are used main effects may not be interpreted. If ordinal data are used the traditional significance tests may produce erroneous results. If non-linear relations exist spurious interaction effects may be found. The problems surrounding moderated regression analysis are largely solved by using alternative significance tests, mean centering and carefully developed (i.e., highly reliable) measurement instruments.

INTRODUCTION: MULTIPLICATIVE INTERACTION

Regression equations of the form $\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_1x_2$ are relatively common in management research (and in the social sciences in general). The term x_1x_2 in the regression equation is commonly labelled the multiplicative interaction term, the influences of x_1 and x_2 are called the main effects. Although the multiplicative term may be derived from a multiplicative model like Vroom's Expectancy Value model (where motivation is a function of expectancy x valence), the interaction term typically is used to test a contingency theory, where the effect of one independent variable on the dependent variable depends on the level of a second independent variable (the moderator).' A contingency theory might for instance claim that the contribution of a certain management technique is dependent on a characteristic of the environment. The contingency theory might, for instance, claim that if the environment is more complex, usage of the balanced scorecard will have a more positive influence on performance than in environments where complexity is low. The theory might even claim that the management technique has an averse influence on performance if complexity is low, and a positive influence when complexity is high. Such an interaction is called disordinal or crossover, whereas an interaction that only indicates a change in level, but not in sign is called ordinal or non-crossover (Lubin, 1961; Aiken & West, 1991).

Mathematically, the link between contingency theory and multiplicative interaction can be interpreted as follows. Starting with a traditional regression equation $\hat{y} = a + b_1x_1 + b_2x_2$, the contingency theory that is bound to be tested posits that b_1—the influence of x_1 on y-depends on the level of x_2 (the moderator): $b_1 = c_1 + d_1x_2$.[2] Substitute b_1 in the traditional regression equation and next expand and reorder it: $\hat{y} = a + (c_1 + d_1x_2)x_1 + b_2x_2 = a + c_1x_1 + b_2x_2 + d_1x_1x_2$.[3] So in order to test for multiplicative interaction y is not only regressed on x_1 and x_2, but also on a third variable that is calculated by multiplying x_1 and x_2. The appropriateness of the contingency model is assessed by comparing the R^2 of the regression including the multiplicative term with that of a regression that contains only both main effects. If R^2 increases significantly after addition of the multiplicative term this is considered evidence in favor of the contingency theory.[4]

This introduction has been used to introduce moderated regression/regression with a multiplicative interaction. In the remainder of this paper some problems and pitfalls associated with this approach will be discussed. In the management literature (with a possible exception of the problems associated with the employment of interval data) the majority of these problems have not been given the attention they deserve. Negligence of these problems, may lead to faulty outcomes of moderated regression analyses, however.

In the next section the effects of measurement level will be discussed and it will be indicated that when the data used for the analysis are measured at ordinal or interval, rather than ratio level, interpretation of the outcomes of the analysis may become problematic. The next section discusses the influence of measurement error. If data are unreliable the results of the analysis may be affected. Next it will be indicated that spurious interaction effects may be found if moderated regression analysis is used on data that in reality contain a non-linear (polynomial) relation between one of the dependent variables and the independent variable. Two relatively minor issues-the exclusion of main effects from the regression equation and standardization— will be discussed in the 'concluding remarks'. The findings of this paper and some practical guidelines for carrying out and interpreting moderated regression will be summarized in a final section.

MEASUREMENT LEVEL

Interval Data

The most commonly acknowledged problem with moderated regression analysis is the Southwood (1978) effect: if data are measured at the interval level, only the regression parameter of the highest order interaction term may be interpreted; only if ratio data are available, meaning can be assigned to all regression parameters. The logic behind this argument is easily made clear. The difference between ratio and interval measurement is that interval measures have an arbitrary origin. The traditional example is the difference between degrees Celsius and Kelvin.

The latter measure is a ratio measure, as its origin is the absolute zero point. Consequently, ratios of ratio measures do make sense, but ratios of interval measures do not. If we compare the values 2 and 4 on a ratio measure like revenue, we can say 4 is twice 2; if the values 2 and 4 are for instance answers on a Likert type question, it would be rather arguable to make such a claim.

Algebraically, an interval measure can be expressed as $x = c + \xi$,[5] where x is the interval measure, ξ is the underlying ratio number and c is an arbitrary but fixed constant. The Southwood effect is easily illustrated by substitution of the xs in the regression equation by this representation; this results in the following expression:

$$\hat{y} = a + b_1(c_1 + \xi_1) + b_2(c_2 + \xi_2) + b_3(c_1 + \xi_1)(c_2 + \xi_2) \quad (1)$$

$$= (a + b_1c_1 + b_2c_2 + b_3c_1c_2) + (b_1 + b_3c_2)\xi_1 + (b_2 + b_3c_1)\xi_2 + b_3\xi_1\xi_2$$

$$= a^* + b_1^*\xi_1 + b_2^*\xi_2 + b_3\xi_1\xi_2 \quad (2)$$

The arbitrary constants (c_i) that are a consequence of the fact that the data are measured at interval rather than ratio level enter the expressions for all (starred) regression parameters with a single exception: the parameter (b_3) of the interaction term. By choosing an 'appropriate' zero point, the other parameters can even be made equal to zero. In other words: if the dependent variables are transformed by adding an arbitrary constant to them, which is allowed if they are measured at the interval level, the regression parameters for the main effect change. It would be hard to assign any meaning to such an inconsistent parameter: the main effects cannot be interpreted. Consequently, only the regression parameter for the interaction term (and its significance) may be interpreted. Of course one would like to be able to do something useful with the other regression parameters as well, but given this representation, at least it is possible to assess the significance of the contingency theory.

The interpretation of the words 'may (not) be interpreted' in the previous paragraph deserves some further attention. The lower order parameters may not be interpreted because they are a function of some arbitrary constants and consequently it cannot be claimed that the findings are universally applicable. However, if two studies use the same measurement instruments, take a large random sample from the same population, and code their findings identically, the constants are identical and consequently the regression parameters may be compared. In fact, this amounts to saying that if a range for the ξs is defined, regression parameters have meaning within this range.

An alternative, but related, approach—later on we will see that this approach has some additional advantageous properties which make its application generally desirable in moderated regression analysis—is to use mean centered data. Before estimating the regression parameters ***and*** before calculating the interaction term (i.e., before x_1 and x_2 are multiplied in order to get the variable x_1x_2), the mean of the variable is subtracted from all observations. Now all regression parameters can be interpreted: they represent the average main effects for all observations or equivalently: b_1 represents the effect of x_1 on y at the mean of x_2 and b_2 represents the effect of x_2 on y at the mean of x_1 (Aiken & West, 1991).[6]

Ordinal Data

The previous section discussed the consequences of measurement at the interval rather than the ratio level. With the introduction of arbitrary constants which are (by definition) associated with the use of interval data, some ability to interpret regression parameters is lost. Fortunately, it is still possible to assess the significance of the multiplicative effect in order to test the contingency theory. Unfortunately, the majority of data used in business research is not measured at the interval level. Typically, Likert scales are used and strictly speaking such data are only measured at an ordinal level. Only the rank of a score may be interpreted; the difference between scores is not necessarily constant. E.g., the difference in the phenomenon of interest that is needed to cause a respondent to mark 5 instead of 4 on a 5-point Likert scale may be larger than the difference that causes a shift from 2 to 3. The problem in this case is that the relation between the true and the observed scores is not known, we only know that one is obtained from the other by applying some strictly increasing function. By definition, our data are interval measures if and only if this function is linear.

If the relation between the predictors (xs) used in the analysis and the phenomenon of interest is not linear the model presented in equation 1 is no longer valid. This problem can also be formulated from the opposite direction: if measurement level is only ordinal, each order-conserving transformation is admissible, and hypotheses tests should be consistent across such transformations. Busemeyer & Jones (1983) present an example where the relation between indicator and phenomenon is approximated by a quadratic polynomial (i.e., $\xi = c + d_1x + d_2x^2$). This polynomial and its domain are chosen in such a way that its inverse exists and this inverse is a function.[7] Expansion of the equation $\hat{y} = a + b_1\xi_1 + b_2\xi_2 + b_3\xi_1\xi_2$ in which the phenomena under investigation are replaced by their indicators is tedious, but straightforward. Finally this results in the following equation

$$\hat{y} = b'_0 + b'_1x_1 + b'_2x_2 + b'_3x_1^2 + b'_4x_2^2 + b'_5x_1x_2 + b'_6x_1^2x_2 + b'_7x_1x_2^2 + b'_8x_1^2x_2^2 \tag{3}$$

In which all b's are a function of the regression constant a, the original regression parameters b and the parameters of the measurement model d. Testing the significance of the interaction term is slightly complicated as all b'-terms are mixed. Fortunately, a significance test is still possible by comparing the R^2 of equation 3 with that of the expanded form of $\hat{y} = a + b_1\xi_1 + b_2\xi_2$, which turns out to be:

$$\hat{y} = b'_0 + b'_1x_1 + b'_2x_2 + b'_3x_1^2 + b'_4x_2^2$$

In other words, testing whether $b'_5 = b'_6 = b'_7 = b'_8 = 0$ is equivalent to assessing the significance of the interaction term (Busemeyer & Jones, 1983).[8] Using the traditional regression equation with only the linear x linear interaction term will produce erroneous results.

Busemeyer & Jones (1983) show a simple example in which invalid conclusions will be drawn if a linear measurement model is inappropriately assumed. If the measurement is quadratic and two variables are correlated with each other (which is not unlikely) a spurious significant interaction term may occur: the interaction term may be significant when no interaction effect of the true variables exists and consequently one may falsely conclude that an interaction is significant (which will be considered evidence in favor of the contingency theory under investigation). Suppose the following equation is appropriate: $\hat{y} = \boldsymbol{a} + b_1\xi_1$. In other words: the only true effect that exists is a linear influence of ξ_1 on y. Now take $\xi_1 = x_1^2$ (in other words a simple quadratic measurement model is assumed) and assume that x_1 and x_2 are correlated. Mathematically this correlation is represented as follows: $x_1 = c + b_2x_2 + e$, where e is random error not correlated with the residual variation in x_1 and c is some arbitrary constant. Now, the regression equation can be rewritten as follows: $\hat{y} = a + b_1x_1^2 = a + b_1x_1(c + b_2x_2 + e) = \boldsymbol{a} + b_1cx_1 + b_1b_2x_1x_2 + b_1x_1e$ and since neither b_1 nor b_2 equals zero a significant interaction term will be found, whereas the true model only contains a direct effect of ξ_1 on y.[9]

Unfortunately, the problems associated with the employment of ordinal data are not easily solved (unless the measurement model is known). A theoretical solution would be to use Taylor series of higher power than the quadratic polynomial presented above to approximate the measurement model. In this case the increase in R^2 that results from extending the expanded main effects only model to the expanded moderated regression model can be used to assess significance of the interaction term. This comparison is similar to that presented for the quadratic measurement model presented earlier. The loss in power would prohibit this approach, however. The only available escape is the hope that a linear or quadratic measurement model resembles the true measurement model closely enough to prohibit the researcher from drawing erroneous conclusions. In that case the traditional significance test (assuming a linear measurement model is appropriate) or the approach presented earlier in this section (assuming a quadratic measurement model is appropriate) can be used for significance testing. This is a rather meager solution, however.

MEASUREMENT ERROR

Until now, this paper only investigated the problems associated with measurement level. Measurement error also deserves attention when testing contingency theory using moderated regression analysis. The traditional psychometric error model is simple enough to allow easy investigation of the effect of measurement errors on parameter estimates for multiplicative models. Each measure is supposed to exist of two components: the true score (ξ) and a measurement error (E): $x = \xi + \epsilon$ (Kerlinger, 1973). As both d ependent and independent variables may be measured with error, the shorthand $\hat{y}$ gets confusing and consequently the error of the regression equation u will be included in the equation explicitly and the letter η will be used to indicate the true score for which y serves as an indicator. This leads to the following true score formulation of the moderated regression model.

$$
\begin{aligned}
\eta - \epsilon_\eta &= a + b_1(\xi_1 - \epsilon_{\xi_1}) + b_2(\xi_2 - \epsilon_{\xi_2}) + b_3(\xi_1 - \epsilon_{\xi_1})(\xi_2 - \epsilon_{\xi_2}) + u \\
&= a + b_1\xi_1 - b_1\epsilon_{\xi_1} + b_2\xi_2 - b_2\epsilon_{\xi_2} + b_3\xi_1\xi_2 - b_3\xi_1\epsilon_{\xi_2} - b_3\xi_2\epsilon_{\xi_1} + b_3\epsilon_{\xi_1}\epsilon_{\xi_2} + u \\
\eta &= a + b_1\xi_1 + b_2\xi_2 + b_3\xi_1\xi_2 + (\epsilon_\eta - b_1\epsilon_{\xi_1} - b_2\epsilon_{\xi_2} - b_3(\xi_1\epsilon_{\xi_2} + \xi_2\epsilon_{\xi_1} - \epsilon_{\xi_1}\epsilon_{\xi_2}) + u)
\end{aligned}
$$

As is clear from the final equation, the error term of the regression is correlated with the independent variables and least squares estimates of the regression parameters *(a, b,, b,,* and b_3) will consequently be biased and inconsistent. The increase in R^2 that results from the change from an additive to a multiplicative model, and on which the test for the appropriateness of a contingency theory is based, will be 'severely attenuated by even moderate measurement error associated with the predictors' (Busemeyer & Jones, 1983, p. 556). According to an analysis presented by these authors the attenuation increases when (1) the number of independent variables in the interaction term increases (i.e., three-way interactions[10] are more strongly attenuated than two-way interactions) and/or (2) the correlation between the independent variables in the interaction term decreases.

Aiken & West (1991) carry out Monte Carlo simulations to assess the importance of the attenuation effect. They find that if reliability of the individual independent variables is 0.8 the variance accounted for by the interaction term decreases by about 50%, 'When reliabilities are .70, the reliabilities accounted for by the interaction is only 33% to 50% of that accounted for when reliabilities are 1.00' (p. 163; the exact percentages are presented in table 8.4 of Aiken & West (1991) and depend on correlation between the independent variables and the level of variance accounted for by the first order terms (main effects)).

Kenny & Judd (1984) suggest a solution for this latter problem by explicitly including measurement in the estimation procedure. They present a structural equation model with latent variables ('LISREL-model'). Unfortunately, even for the case where only two independent variables are used and each is measured by only two manifest variables ('questions'), this model is exceedingly complex, LISREL's maximum likelihood estimation procedure is no longer applicable and non-linear constraints need to be included in the model. For most management studies the structural equation modeling approach seems to be too ambitious a goal to aim for.

Alternative solutions have been proposed, but they require uncorrelated measurement errors,[11] have a tendency for overcorrection and may prohibit significance testing of the parameters (Aiken & 'West, 1991). A more realistic aim is the application of well validated measurement instruments with high reliability ($r \gg 0.80$) and to aim for low correlation between the variables included in interaction terms. The latter property may be reached by the employment of mean centered data.

POLYNOMIAL EFFECTS

Another problem associated with moderated regression analysis is to distinguish interaction effects from polynomial (e.g., quadratic) effects. If the true model is quadratic and the independent variables used in the traditional multiplicative regression equation are correlated a significant, but spurious, interaction term may be found. In other words, false support for a contingency theory may be found if non-linear effects are present. Mathematically, this problem may be illustrated as follows. In the true model $\hat{y} = a + b_1\xi_1 + b_2\xi_2^2$. Now simply take $x_1 = \xi_1$ and assume that x_1 is correlated with a second variable x_2 and consequently $x_1 = c + x_2 + e$. Now $\hat{y} = a + b_1x_1 + b_2x_1^2 = a + b_1x_1 + b_2x_1(c + x_2 + e) = a + b_1x_1 + b_2cx_1 + b_2x_1x_2 + b_2ex_1 = a + (b_1 + b_2(c + e))x_1 + b_2x_1x_2$, and since b_2 does not equal zero support for the existence of an interaction term will be found even if no interaction effect exists.

In order to determine whether a true interaction effect exists, according to Lubinski & Humphreys (1990) not only the model containing the interaction term ($\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_1x_2$) but also both models containing the quadratic effect ($\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_1^2$ and $\hat{y} = a + b_1x_1 + b_2x_2 + b_3x_2^2$) should be estimated; next the model with the highest R^2 should be selected. In a comment on the Lubinski & Humphreys paper, Shepperd (1991) indicates that this advice may itself lead to spurious results as it is data driven and leads to capitalization on chance. MacCallum & Mar (1995) introduce the concept of measurement error and multicollinearity into this discussion. They refer to Busemeyer & Jones (1983) who discussed the influence of the reliability of the individual independent variables and the correlation between the independent variables on the reliability of the interaction term. Busemeyer & Jones indicated that the reliability of the interaction term (i.e., x_1x_2) is positively associated with the correlation between the independent variables. If the variables are uncorrelated, the reliability of their product equals the product of their reliabilities. If they are correlated reliability increases. The reliability of a squared variable (i.e., x^2) simply is the square of the reliability of that variable. Consequently, for correlated variables the reliability of the interaction term will tend to be higher than the reliability of the squared term.[12] As a consequence of this difference in reliability, the model including the interaction term will tend to be preferred over the quadratic model and consequently support for contingency models will be found too often.

MacCallum & Mar (1995) carry out a simulation study to investigate how severe the theoretical problems mentioned above are in practice. If the true model is multiplicative, the R^2-based selection procedure tends to make the correct choice by providing support for the multiplicative rather than the quadratic model. If (1) multicollinearity between the independent variables is severe, (2) the effect size is small, (3) reliability of the independent variables is low or (4) the number of observations is small, the results get somewhat worse, but even in the worst case (a low effect size, item reliabilities of 0.7, 75 observations and a correlation of 0.9 between the true scores of the independent variables) the correct classification is made in more than 50% of the simulations. For the opposite situation, the results are less positive. If the true model is quadratic, the multiplicative model tends to be selected rather often. In particular reliability of the independent variables and multicollinearity between these variables appears to be an important determinant of the number of times an incorrect model is selected.

The results of MacCallum & Mar (1995) can be used to derive the following rule of thumb: item reliabilities should be above 0.8 and the correlation between true scores[13] of the independent variables should be below 0.7 if the number of observations is at least 300. If the number of observations is lower, reliabilities of the independent variables should be somewhat higher and in particular multicollinearity should be lower (say at most 0.6). In general, less is to be gained from an increase in the number of observations, than from an increase in reliability or a decrease in multicollinearity. If the requirements are met the best way to compare a multiplicative and a quadratic mode1 is to select the one with the highest R^2. The difference in R^2 between the quadratic and multiplicative model need not to be significant (and often will not be significant).

An alternative solution, of course, is to include both the quadratic effects and the interactive effect in the regression equation. However, this approach may result in loss of power (Ganzach, 1998). Simulation results of Ganzach (1998) show that adding a quadratic term if the true model does not include a quadratic effect does result in the expected loss of power. However, 'for the parameters typically encountered in management research, the increase in type II error associated with adding the quadratic terms is not large, and as the number of observations rises, it becomes quite minimal' (p. 619). The loss in power gets stronger if multicollinearity between the independent variables is higher. If the true model contains both quadratic effects and an interaction term, addition of a quadratic term may both result in an increase (if the sign of the quadratic effects is opposite to that of the interaction effect) and a decrease (if quadratic effects and interaction effects have the same sign) of the probability that the interaction effect is found (Ganzach, 1998). Whether the signs are opposite or not, inclusion of the quadratic terms is warranted, as the parameter estimates and the significance of the interaction term will be biased otherwise. The most prudent advice apparently is to include the quadratic terms unless the number of observations is too small (say below 150). If the quadratic terms are not included results of an analysis including the terms should at least be presented next to the results of the analysis.

SOME CONCLUDING REMARKS

In this section two remaining issues that got some attention in the literature, but are relatively independent of the remainder of the issues raised in this paper, will briefly be discussed. First some remarks will be made about the suggestion to omit main effects from the regression and to estimate a regression equation that does only contain the interaction term. Next a relatively minor issue will be discussed: how does standardization affect the analysis of interaction effects?

Stone & Hollenbeck (1984) criticized an issue they observed in the literature: some researchers apparently choose to estimate the equation $\hat{y} = a + b_3^* x_1 x_2$ instead of the full interactive regression equation discussed earlier in this paper. Apparently the logic followed is that if the interaction term is correlated with the variables it is calculated from, the power of the significance test for the interaction term is reduced and consequently it would be better to eliminate the main effects completely. Eliminating the main effects does indeed increase the chance that b_3^* is significant. However, Stone & Hollenbeck (1984) argue and demonstrate that this is not desirable. The resulting regression overestimates the amount of variance accounted for by the interaction term and consequently is too likely to find a significant interaction when it does not exist. Main effects should be included in the regression equation, unless very strong theoretical evidence exists that they exactly equal zero (Cohen & Cohen, 1983) and even in the latter case it would probably be relatively harmless to include them.[14] Of course mean centering can be used to reduce the correlation between the interaction term and the independent variables it is calculated from. In addition, it should be noted that if convincing evidence for the existence of a purely multiplicative model exists, the application of appropriate data transformations (by taking logarithms) may be preferred over the use of moderated regression analysis.

Another minor issue is the question whether differences in correlation coefficients across groups may be considered evidence for the existence of an interaction effect. If the correlation between experience and salary is significantly lower for men, than for women, does this provide evidence for a contingency theory that claims that the relation between salary and experience is moderated by gender? Although the terminology used by Arnold has been criticized by Stone & Hollenbeck (1984), the logic of his argument (Arnold, 1982, 1984) is rather clear: it does not. Different correlations do not provide evidence of a different influence, they may also be caused by the fact that standard deviations for the variables are not equal for both groups. Typically the claim of a contingency theory will be that the contingency variable (gender in the example presented above) affects the influence of a second variable (experience in this case), rather than claiming that the predictability of the dependent variable will be influenced. Consequently, only moderated regression is appropriate, correlation analysis is not.[15]

Investigation of correlation coefficients rather than regression parameters is a form of standardization. Standardization itself deserves some additional attention. Not only the analysis of correlation coefficients, but even the investigation of standardized solutions to a regression analysis should be treated with care. Generally speaking, if the regression model does contain an interaction term, the standardized solution produced by standard statistical software is not useful: a standardized solution should be produced 'manually' by first standardizing the dependent and independent variables, next calculating a 'standardized' interaction term by multiplying the standardized independent variables. On the dataset thus obtained a traditional moderated regression analysis can be performed and the ***unstandardized results*** of this analysis provide the ***standardized*** solution (Aiken & West, 1991). The critical aspect in this approach is that it ensures that the interaction term is the product of the independent variables, which is not the case if the interaction term itself is standardized. As a consequence of this approach the standardized solution has the unusual feature that it does include a constant (i.e., the a of the regression equation which is the standardized solution, does not necessarily equal zero).

CONCLUSION

The inevitable conclusion of this paper is that moderated regression analysis is a technique draught with potential pitfalls. This does not make the technique useless or superfluous, however. In some case-for instance when analyzing truly multiplicative models-the application of alternative approaches like data transformations may be a solution, but for contingency theory moderated regression analysis remains the most appropriate technique. Although moderated regression analysis cannot always be defended on theoretical statistical grounds, mathematical analysis of its properties and the simulation studies discussed earlier in this paper suggests that the results are rather robust if the proper precautions are taken.

A first general advice is to use mean centered data. First this may solve potential computational problems which otherwise might be caused by multicollinearity. Second and more important, after mean centering the main effects of the moderated regression become interpretable: even if the data have not been measured at the ratio but only at interval level they represent the average main effects of the independent variables for all observations.

A second general advice—which almost is a truism-is to use reliable measurement instruments. Moderated regression analysis assumes the absence of measurement error. Measurement error can be modeled using alternative techniques, but given the complexity of such modeling and the number of observations required to estimate such models, this seems to be a bridge too far for the majority of business research. Fortunately, the problems of unreliable measurement seem to be rather minor if reliability is high (reliabilities of at least 0.8 and preferable of 0.9 or more are desired).

A third precaution against drawing erroneous conclusions is to include terms representing plausible alternative models whenever possible. Especially main effects should never be omitted from the regression equation. Inclusion of quadratic effects may cost too much power if the number of observations is low (say below 150), but if the existence of such effects is conceivable at least the results of an (additional) analysis including the quadratic term should be reported.

NOTES

[1]Although one can dispute the question whether multiplicative and contingency models are conceptually identical, at least mathematically an interaction term does not discriminate between both concepts.

[2]Note that an implicit assumption is introduced in this way. The contingency theory just claims that the effect of x_1 on y depends on the level of x_2. Testing of this equation by regression with multiplicative interaction implies that the effect of x_1 on ***y linearly*** depends on x_2.

[3]This substitution is symmetric. Both the claim that b_1 linearly depends on x_2, and the claim that b_2 linearly depends on x_1, and the claim that both b_1 and b_2 depend on x_2 and x_1 respectively will result in the multiplicative equation.

[4]The moderator variable does not need to be continuous, dichotomous variables may be used as well, and multiple interactions with dummy variables can be used to model the moderating influence of a nominal variable (Cohen & Cohen, 1983).

[5]In this paper Greek characters are used to indicate a variables true score and Latin characters are used to indicate the observed score. The choice of the variable names is based on the names traditionally used in LISREL modeling (Jöreskog & Sörbom, 1989).

[6]Algebraically it is easily seen that the problem of arbitrary constants that comes along with the use of interval data disappears in this case. A change in this constant causes an identical change in the mean which annihilates the effect observed earlier.

[7]An example of such a function is the application of the following formula to a 5-point Likert scale $\xi = -\frac{1}{2} + \frac{1}{2}x^2$, in other words a true score of zero results in answer one, a true score of 1.5 in answer two, a true score of four in answer three etc. Such increases in true scores necessary before a higher answer category is selected do not seem unlikely to exist at all.

[8]Note that, apart from a loss of power, nothing is wrong with using this test if the measurement model happens to be linear

[9]The question whether the assumptions of Busemeyer & Jones are realistic is not in order here. If

data are measured at the ordinal level, hypotheses tests should be consistent across order preserving transformations. Even a single counterexample like the one presented here proves that this is not the case.

‘*Three-way interactions are not discussed in this paper. A three-way interaction includes a term like $x_1x_2x_3$. Three-way interactions reflect the expectation that the extent to which an independent variable influences the influence of a second independent variable on the dependent variable depends on a third independent variable. In other words three-way interactions can be used to test theories that claim that the form of a contingency theory is contingent upon some other characteristic.

[11] Correlated errors are highly likely if data suffer from common method bias, for instance because all data are collected using a questionnaire survey.

[12]Strictly speaking these results have only been rigorously demonstrated for centered variables. However, intuition and simulation studies suggest that the results are also applicable to uncentered variables (MacCallum & Mar, 1995). Intuitively the higher reliability of the interaction term may be made plausible by looking at this term as a special kind of multi-item measure. If the same variable is used twice (that is in the quadratic term), no additional information is used in calculating the multiplicative term. If two different variables are used, the interaction term will contain additional information; it is unlikely that the variables have perfectly correlated measurement errors.

[13]Correlation between the true scores of course is not directly available in empirical studies. However, it may be estimated from the reliabilities of the independent variables and their observed correlation. Call the correlation between the observed scores $r_{x_1x_2}$, the correlation between true scores $r_{\xi_1\xi_2}$, and the reliability of a variable $r_{x_ix_i}$, then $r_{\xi_1\xi_2} = \frac{r_{x_1x_2}}{\sqrt{r_{x_1x_1}}\sqrt{r_{x_2x_2}}}$ (Nunnally, 1967; MacCallum & Mar, 1995). If for instance both independent variables have a reliability of 0.8 and the correlation between both variables is 0.5, then the best estimate of the correlation between the true scores of the independent variables is $\frac{0.5}{\sqrt{0.8}\sqrt{0.8}} = 0.625$.

[14]Results of a Monte Carlo simulation carried out by the author of this paper, tend to confirm this intuition. If the number of observations is at least 50, reliability is above 0.7 and R^2 is at least 0.10 (and preferably above 0.25) the inclusion of main effects, whereas the data a generated from a multiplicative

(‘interaction term only’) model, are relatively harmless.

[15]Note that the logic above is formulated in terms of contingency theory. In personnel psychology situations may be encountered where predictability itself is an issue. E.g., when assessing the validity of management tests or similar personnel selection instruments. Such ‘moderators’ have been called ‘predictors of predictability’ (Zedeck, 1971; Zedeck et al., 1971). They determine the extent to which a function is applicable to a group of subjects, rather than the existence of different functions for different groups.

REFERENCES

Aiken, L. S., & West, S. G., 1991. ***Multiple regression: testing and interpreting interactions.*** Sage. Second printing, 1992.

Arnold, H. J., 1982. Moderator variables: a clarification of conceptual, analytic, and psychometric issues. ***Organizational Behavior and Human Performance, 29:*** 143-174.

Arnold, H. J., 1984. Testing moderator variable hypotheses: a reply to Stone and Hollenbeck. ***Organisational Behavior and Human Performance, 34: 2 14-224.***

Busemeyer, J. R., & Jones, L. E., 1983. Analysis of multiplicative combination rules when the causal variables are measured with error. ***Psychological Bulletin, 93 (3): 549-562.***

Cohen, J., & Cohen, P., 1983. ***Applied multiple regression/correlation analysis for the behavioral sciences.*** Lawrence Erlbaum Associates, Hillsdale, New Jersey; London, 2nd edition.

Ganzach, Y., 1998. Nonlinearity, multicollinearity and the probability of type II error in detection interaction. ***Journal of Management, 24 (5):*** 615-622.

Joreskog, K. G., & Sörbom, D., 1989. LISREL7***: a guide to the program and applications.*** SPSS Inc., 2nd edition.

Kenny, D. A., & Judd, C. M., 1984. Estimating the nonlinear and interactive effects of latent variables. ***Psychological Bulletin, 96*** (1): 201-210.

Kerlinger, F., 1973. ***Foundations of behavioral research.*** Holt, Rinehart, and Winston, 2nd edition.

Lubin, A., 1961. The interpretation of significant interaction. ***Educational and Psychological Measurement,*** 21 (4): 807-817.

Lubinski, D., & Humphreys, L. G., 1990. Assessing spurious 'moderator effects': illustrated substantively with the hypothesized ('synergistic') relation between spatial and mathematical ability. ***Psychological Bulletin,*** 107 (3): 385-393.

MacCallum, R. C., & Mar, C. M., 1995. Distinguishing between moderator and quadratic effects in multiple regression. ***Psychological Bulletin,*** 118 (3): 405–421.

Nunnally, J. C., 1967. ***Psychometric theory.*** McGraw-Hill.

Shepperd, J. A., 1991. Cautions in assessing spurious 'moderator effects'. ***Psychological Bulletin,*** 110 (2): 315-317.

Southwood, K. E., 1978. Substantive theory and statistical interaction. ***American Journal of Sociology:*** 1154-1203.

Stone, E. F., & Hollenbeck, J. R., 1984. Some issues associated with the use of moderated regression. ***Organizational Behavior and Human Performance, 34:*** 195–2 13.

Zedeck, S., 197 1. Problems with the use of 'moderator' variables. ***Psychological Bulletin,*** 76 (4): 295-310.

Zedeck, S., Cranny, C. J., Vale, C. A., & Smith, P. C., 1971. Comparison of 'joint moderators' in three prediction techniques. ***Journal of Applied Psychology, 55 (2): 234-240.***